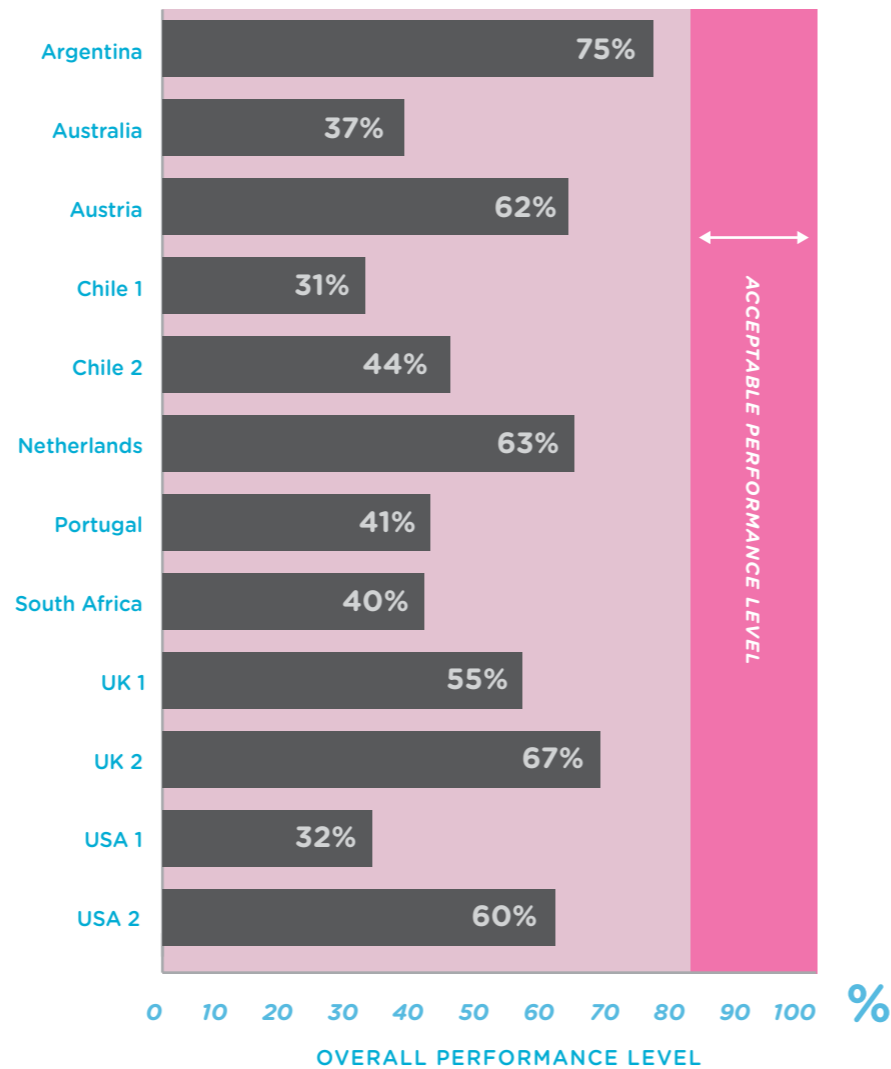


Credit card statements

BASELINE MEASUREMENT 2009



Summary report

DAVID SLESS





e-papers

Credit Card Statement

Baseline measurement 2009

David Sless

© 2014 COMMUNICATION RESEARCH INSTITUTE. ALL RIGHTS RESERVED.

WEBSITE [HTTP://COMMUNICATION.ORG.AU](http://communication.org.au)

PHONE +61 (0)3 9005 5903

ADDRESS FITZROY NORTH • MELBOURNE • AUSTRALIA • 3068

ABN 74 120 911 131

Contents

What are communication baseline measurements?	3
Background to this study	4
Method	6
Results	11
Discussion	12
Conclusion	18
References	20

Figures

Figure 1: Typical cumulative data on design faults	10
Figure 2: Participating countries	11
Figure 3: Summary performance across all CCS	12
Figure 4: Statement performance	13
Figure 5: Task performance	15

Tables

Table 1 : The investigators	5
Table 2: CCS Performance requirements	7
Table 5: Participant comments	17

What are communication baseline measurements?

30 years ago the Communication Research Institute (CRI) began baseline measurement studies of communication practices used by business and government in their communication with the public.

Communication baseline measurements quantify the number and types of faults in a design to see how far short of an acceptable performance level the design falls. Baseline measurements also provide a great deal of qualitative data on the causes of the failure.

The types of communication CRI studies include voice systems, forms, legal documents, bills, letters, product labelling, consumer instructions, and websites—the stuff of ordinary life that originates from business and government and makes up a large part of the daily communication between organisations and the public.

In the mid-1990s CRI undertook communication baseline measurements in Australia of banking websites ⁽¹⁾, medicine labelling ⁽²⁾, financial services guides ⁽³⁾, government and business forms ⁽⁴⁾, utility bills ⁽⁵⁾, Consumer Medicines Information (CMI) ⁽⁶⁾, and many other types of designed information.

Detailed data from these studies are used by CRI Fellows to help our Corporate Members. Wherever possible, we avoid drawing attention to specific institutions. We have no interest in ‘naming and shaming’. Rather, our interest is in drawing attention to current public communication practices, in order to encourage the whole of industry and government to improve their practices in the future. The data we provide establish the communication baseline measurements against which we can measure their future improved practices.

Background to this study

Credit cards are used by billions of people all over the world, and in the wake of that use come the inevitable credit card statements (CCSs), the bills to be paid. These statements are at the heart of the communication between credit providers and their customers. They are the main instrument through which customers conduct their business with the credit providers.

This communication baseline measurement study is the first international study of its kind. We chose for this first international study a document—the credit card statement (CCS)—that is relatively widespread around the world, and which has attracted considerable public attention recently.

All those involved in the baseline measurements activities were volunteers. Twelve CRI Fellows and Subscribers from Australia, Austria, Chile, Netherlands, Portugal, South Africa, UK and USA gave freely of their time as investigators on the project, and they recruited other volunteers to participate in the study. The whole project was managed by CRI in Melbourne.

Table 1 : The investigators

Consuelo Amenabar, *Chile*

Thomas Bohm, *UK*

Veronika Egger, *Austria*

Sandra Fisher-Martins, *Portugal*

Martin Gallo, *Argentina*

Frances Gordon, *South Africa*

Claudine Jaenichen, *USA*

Judith Moldenhauer, *USA*

Jane Teather, *UK*

Alexander Tyers, *Australia*

Karel van der Waarde, *Belgium*

Carola Zurob, *Chile*

With the help of our volunteers, we collected a convenience sample of CCSs from around the world. These were depersonalised to remove any information that would identify specific card holders.

This is a report of their work: the findings made possible by their collective efforts and collaboration. Both the investigators and the participants have made an important contribution to our field, and we hope they will continue to collaborate with us in the future. Indeed, we hope that their example will encourage many others to join us in our Communication Baseline Measurements program.

Method

Diagnostic testing sessions

The method used in this study is called diagnostic testing⁽⁷⁾. It is conducted in a session involving an investigator and a participant. As its name suggests, the method was developed specifically to help information designers identify and diagnose design faults.

There is now a body of evidence and experience in the use of diagnostic testing that confirms its technical reliability, sensitivity and validity⁽⁸⁾. There are also good research conclusions on the best types and numbers of people that are needed as participants, in order to get useful data⁽⁹⁾.

Like diagnostic tools used in medicine, diagnostic testing in information design is at its most powerful when used in a context where the presence or absence of symptoms of 'pathology' is used as a guide to the most appropriate 'treatment'.

Diagnostic testing sessions are conducted one-on-one in a quiet room. The investigator and each participant singly collaborate in a conversation around the use of a particular document. The investigator makes it clear to the participant at the outset that the purpose of the diagnostic session is to find out through the diagnostic testing what, if anything, is wrong with the document. The investigator asks the participant to undertake a number of tasks with the document, recording what they do and say whilst trying to complete the task. Participants are prompted to talk about what they are doing and any problems they encounter.

The investigator makes three types of quantitative observation with each participant:

- ❖ Can they find the information?
- ❖ Do they have difficulty finding the information?
- ❖ Can they use the information appropriately once they have found it?

Alongside this quantitative data, investigators report their detailed observations on the types of difficulty participants have in finding information, and report the verbatim comments of participants throughout the diagnostic sessions.

Performance requirements and protocol

In this study, using the CCS as our starting point, we developed a set of performance requirements for this type of document. Performance requirements consist of two things: a list of the tasks that we believe people should be able to perform with the document, and an acceptable level at which we expect people to perform those tasks.

Usually, the process of compiling and agreeing to a set of performance requirements involves extensive consultation with all stakeholders. However, in this case, we used our own prior experience with many similar documents. The tasks people might be expected to perform with these documents are of two sorts: tasks that have to be performed on any bill that needs to be paid, and tasks that might have to be performed by anyone using a credit card. Table 1 shows the performance requirements developed for this study.

Table 2: CCS Performance requirements

IDENTIFICATION TASKS	BASIC USAGE TASKS	INTERACTIVE TASKS
Identify what the document is (a credit card statement)	Find and explain the date range covered by the statement	Find and explain how to make a payment*
Find and identify who is providing the statement (company name)	Find and explain the opening balance	Find and explain how to find more information
Identify who the credit card statement is for (name, address, account number)	Find and explain the closing balance	
Find and explain the statement period (i.e. monthly statement, annual statement)	Identify the total of any cash advances for the statement period and the interest rate that applies*	
	Identify the total of any purchases for the statement period and the interest rate that applies	
	Find and explain any interest that has been charged to the account	
	Identify any transaction dates	
	Find and explain any transaction descriptions	
	Find and explain the overall credit limit	
	Find and explain any available credit	
	Find and explain any payments that have been made*	
	Find and explain any payments due (when, how much, any overdue amounts)*	
	Find and explain any terms and conditions*	
	Find and explain how many pages are included in the statement	

The target performance level that we aim for is that any customer using a credit card statement be able to find at least 90% of what they are looking for, and

then appropriately act on 90% of what they find (i.e. demonstrate they can use it appropriately).

These two figures are based on:

- ❖ our information design experience
- ❖ our research findings, which demonstrate that these levels are achievable
- ❖ our extensive consultation with stakeholders from industry, government, and consumer advocates who have agreed to these target performance levels are acceptable.

To provide a headline figure, we multiply the percentage found by the percentage used appropriately. This gives us a target performance level of 81% (90% x 90% = 81%).

This composite figure is a 'headline'; it draws attention to the presence of faults in the design. When the components making up these numbers and their related qualitative data are examined together, a full diagnosis of each fault can be undertaken. These overall figures provided us with a picture across all the tasks participants performed and all the CCS that were tested. (As an example of this at work in practice, see: [Australian Self Medication Industry \(ASMI\). Labelling code of practice for designing usable non-prescription medicine labels for consumers.](#))

The performance requirements listed in Figure 2 were used to develop the test protocol—the list of questions and requests to participants to undertake the tasks specified in the performance requirements—that the investigators used in the diagnostic sessions.

The test protocol is designed to form the basis of the one-on-one conversation between an investigator and a participant. The test protocol in this study needed to work across all credit card statements. It was pilot tested by an investigator in two diagnostic sessions to detect any problems that needed to be resolved before finalising it for use. It was also vetted for individual statement suitability with each investigator prior to commencing the study.

Twelve credit card statements from around the world were provided by our volunteer investigators.

Our investigators followed the diagnostic procedure outlined above, using the same protocol translated into the local language where needed. They collated the data on standardised spreadsheets and returned them to CRI in Melbourne, where we checked them, conferred with each investigator to resolve any queries, and then analysed and aggregated the data.

All personal information about the specific participants at each session remained confidential and were not passed on to our project manager. The performance data collected are presented in the following section.

Diagnostic logic

Conventional thinking suggests that the focus of diagnostic sessions is people, that it is people who are being tested. But if we take that view, then we would be required to offer an explanation of the results in terms of people—not just their actions but their inner cognitive processes as well. While we can observe peoples' actions, we have no access to their inner cognitive processes, and consequently we would be involved in a set of inferences based on current cognitive theory—not the firmest foundation on which to build an explanation of what is happening in this context. Anyway, we do not have to be cognitive scientists, as we are not in the business of changing people. We are in the business of changing designed information.

Also, if the focus is on the people, there is an implied criticism of them: it is the people who are having difficulty using a document and the implication is that it is their fault. Most commonly this leads to the easy argument that if people are having difficulty reading,

they have a 'literacy problem', or a 'financial literacy problem'. These terms are used as little more than a way of excusing poor document design. Thus there is no need to redesign a document, because the problem lies in people's deficiency. Time and time again, our research shows that if there is a 'literacy problem' it is in the organisations producing the documents, not in the people who are the hapless victims of this illiteracy. Blaming the victims does not get to the cause of the problem, nor does it solve it.

The diagnostic logic we follow is to take people's actions with a document as symptoms of the underlying condition of the documents themselves. If a document cannot be used for a particular reasonable purpose, then there is a fault in the document. The pathology is in the document, not the people who try to use it and fail. Moreover, if the document is redesigned so that it can be used successfully, we take this as evidence that it is the document that was sick, not the hapless user, and that it has been cured of its pathological condition.

Sample size and data quality

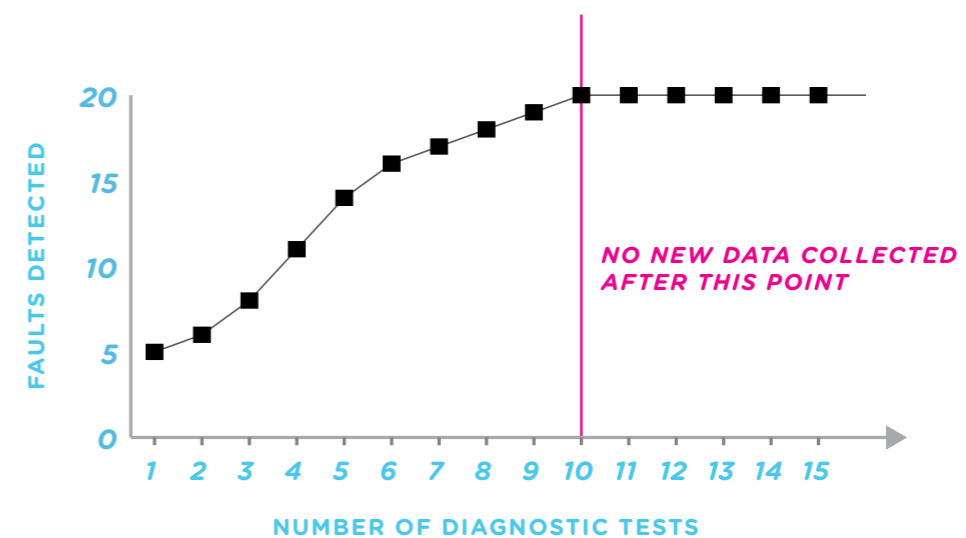
We are often asked: How many people do you test in order to get useful data? The short answer, using the above diagnostic logic, is 'None'. We don't test people, we test the information they try to use. This may seem an odd answer, particularly if you come from a background steeped in social science research methods, but the force of this quick answer lies in the way it directs attention away from the study of people to the study of information. We aim to bring about desirable changes in everyday information, not to bring about change in the people who have to put up with this information.

The longer answer is very much tied to what we are investigating, namely the faults in designed information. The question we ask is subtly inflected by this interest: 'How many diagnostic sessions do we need to conduct in order to identify all the faults in a design? and we answer: As many diagnostic sessions that it takes until we stop collecting any new data about a design's faults.

The cumulative evidence from research and experience suggests that the first six diagnostic one-on-one sessions, each with a different participant, enable the researcher to identify approximately

80% of the faults in a design arising from the tasks participants are asked to perform. After ten such sessions, approximately 100% of these faults have been detected. No new data is collected in the eleventh and subsequent sessions. Figure 2 shows a typical pattern of the cumulative data in such studies.

Figure 1: Typical cumulative data on design faults



Results

A total of 97 diagnostic sessions were conducted with 12 CCSs in 9 countries.

Figure 2: Participating countries



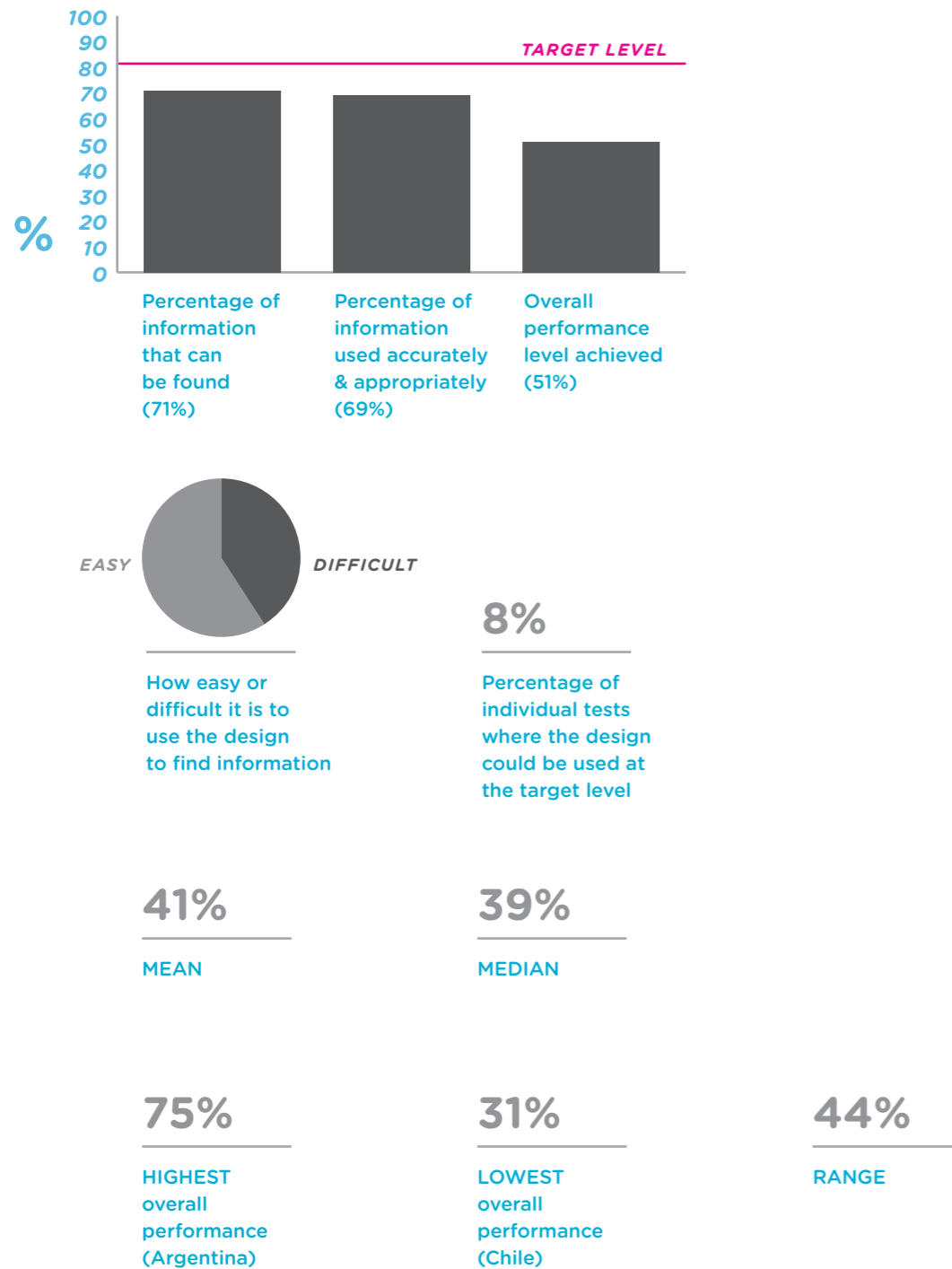
The overall performance level of the CCS

The three types of quantitative observations made by investigators with each participant (could they find the information, did they have difficulty finding the information, could they use the information appropriately once they found it) leads to simplified scoring which provides useful headline figures indicating the overall performance of a document.

But when taken with the investigators' notes from observing participants' actions and writing down participants' verbatim comments, the result is a detailed story rich with data, much of it providing invaluable qualitative insights into the faults and the reasons for them. These data are extremely valuable not only for information designers helping industry improve their designs, but also for regulators to identify key performance indicators to incorporate into regulations to lift the minimum standards of CCS to an acceptable level. In this section we concentrate on the headline figures which are of more interest to the general reader rather than the specialist information designer.

Figure 3 gives a summary of the overall performance across all credit card statements.

Figure 3: Summary performance across all CCS



Discussion

The fact that no statements could be used to perform the basic tasks that are expected of statements indicates that the standard approach to statement design is deeply flawed. In other words, improvements can only be made through a fundamental shift in design thinking and approach.

We were actually surprised by just how poor the performances of the statement designs were, and how universally poor the performances were across all of the designs. Given the simplicity of much of the information provided on a credit card statement, there is a huge scope for improvement.

A CCS is firstly an itemised bill, secondly a detailed record of transactions within the bill, and thirdly, an

account of the business rules applied by the service provider.

Most of the problems stemmed from the inability of participants to use the design to find what they were looking for, as there was little information in the statements to guide participants towards what they were looking for.

Only 8 of the 97 participants could successfully use the statement above the target performance level of 81%.

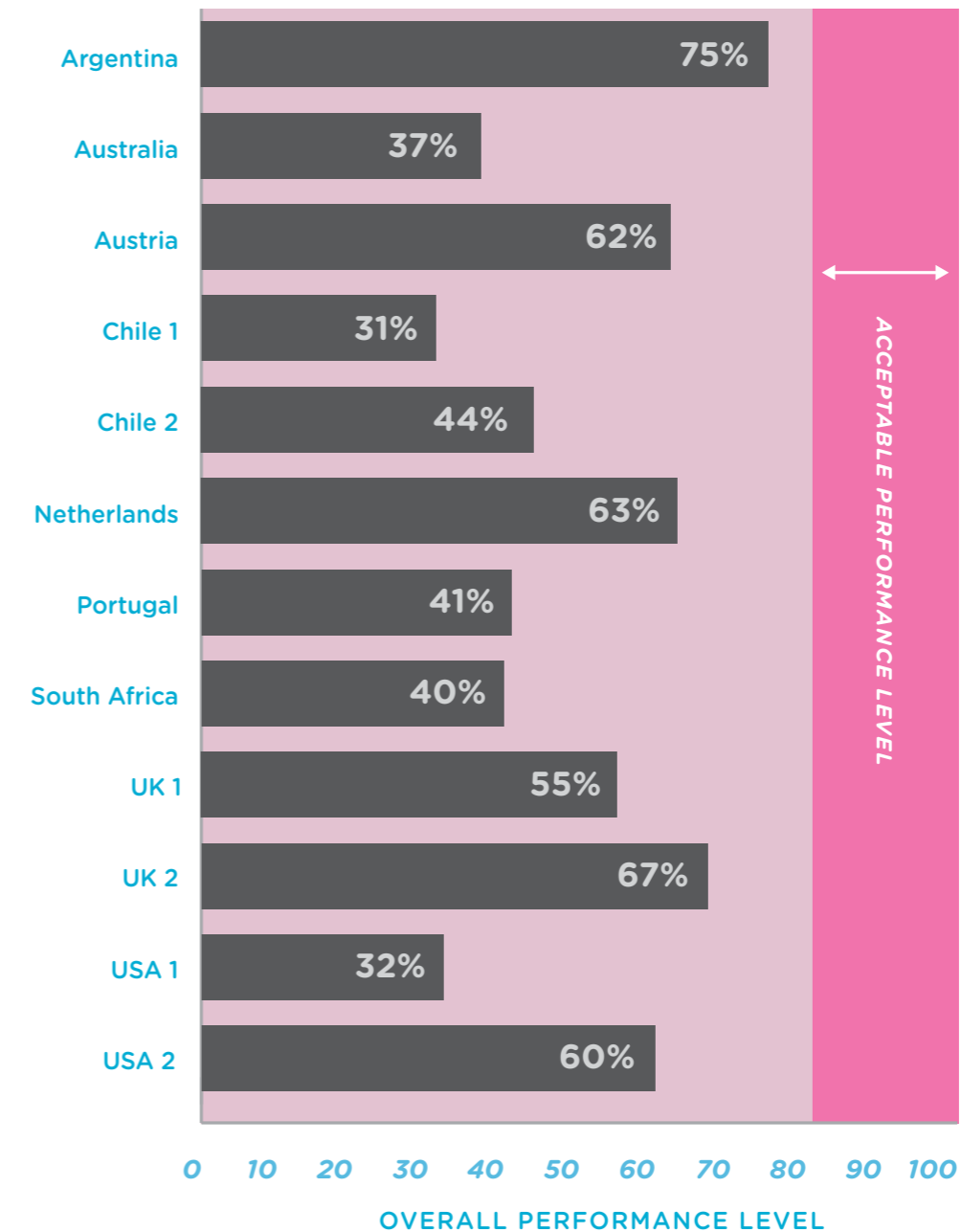
Participants struggled to find what they were looking for 41% of the time, which, given that most of the content appears on a single A4 page, is very poor. Participants could find only 71% of the information they were looking for.

A mere 69% of the information on the statements could be used appropriately. In other words, three out of every ten content items on the statement were unusable.

We expected that at least some of the 11 CCSs studied might achieve an acceptable performance level. But aggregating the data for all tasks performed on each CCS showed that none achieved an acceptable overall performance level of 81%. That is, none of the CCS tested could be successfully used to find 90% of the

information, and when found, successfully used on 90% of occasions. Figure xx shows this aggregated set of results

Figure 4: Statement performance



Performance of each CCS

None of the statements could be used at the overall target performance level, 81%, that was set by our study.

The credit card statements that we tested performed at varying levels—ranging from a low of 31% to a high of 75%—all below the target performance level.

The best performed statement was the Argentinian sample, at an overall level of 75%, followed by the UK (2) sample at 67% and the Netherlands sample at 62%. The worst performances were by the Chilean (1) sample at just 31%, the USA (2) sample at 32% and the Australian sample at 37%.

The Argentinian and Netherlands statements could be used at the target performance level of 81% by 30% of participants in testing; the South African statement was next at 25%; and the Austrian, the second UK and second US statements followed at just 10%. For the statements of the other countries, not one participant reached the target level.

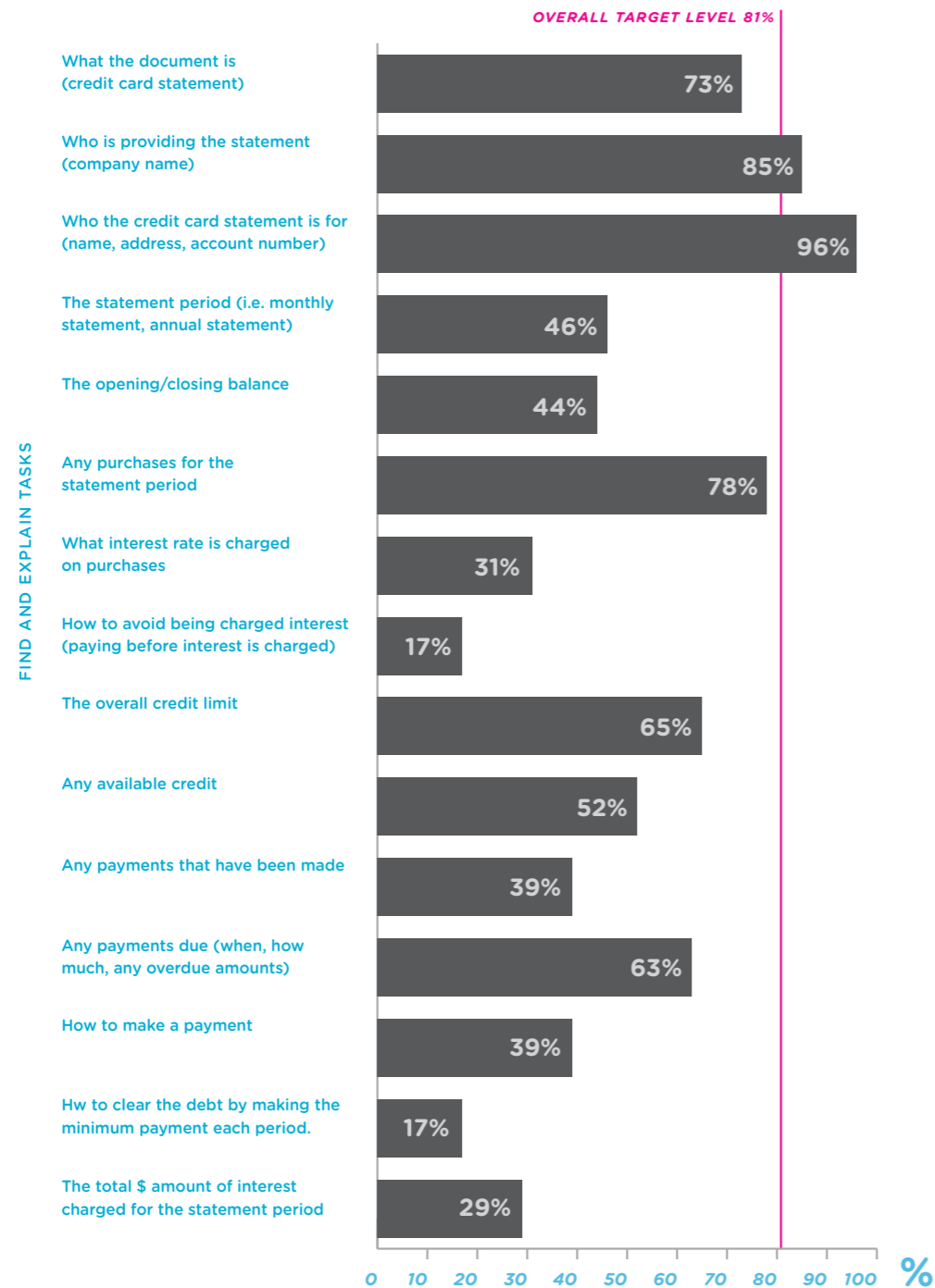
The worst individual test was registered for the Australian credit card sample with an overall performance of just 6%. The second worst was registered by the first Chilean sample, at just 8%. The

best individual performance was 100% registered by one participant using the Austrian sample, who could use the design to find and explain information for every task. The next best performance was by the Argentinian sample, which could be used on one occasion at a level of 94%.

Some statements, such as the Austrian, USA 2, Portuguese, and Chilean 1 samples, lacked the key information needed for customers to gain a basic understanding of their credit card usage and the charges that apply. Others, such as the Australian sample, provided key information (e.g. the amount of interest that had been charged for the statement period) in fine print in obscure and difficult to find locations on the statement.

Performance of tasks

Figure 5: Task performance



The aggregated data for each of the tasks, across all the CSS tested and all diagnostic sessions (Figure 5) show that only 2 tasks out of 14 reached the target performance level:

- ❖ Find and identify who is providing the statement (company name).
- ❖ Find and identify who the credit card statement is for (name, address, account number).

Twelve tasks were below the target performance level, some well below. These under performing tasks ranged from 17% (avoiding interest charges) to 78% (identifying purchases on card for statement period).

When it comes to using the CCS to work out the business rules, the CSS provides little help. On average, the CCS could only be used 29% of the time to work out the interest rate that was being charged. In three cases this was because the information was simply not there to be found. On average, only 17% could work out how to avoid interest payments, and only 17% could use the CSS to work out the consequences of paying the minimum amount due each month.

Only 65% could identify the credit limit; related to this, only 52% could identify how much credit was left and

consequently may not have been able to work out what they had spent.

The most important information that consumers want to know about a bill is how much they have to pay, when they have to pay by, and how to make a payment.

As Figure 6 shows:

- ❖ only 63% could use the CCS to work out what payment was due and when
- ❖ only 39% could use the CCS to work out how to pay.

A few consumers will go to the next level of detail and want to know the way in which the various items on the bill are charged. Here too the CCS presents consumers with a challenge. On average, 78% could use the CCS to work out the items that were being charged for (this percentage would probably be higher if they were looking at their own transactions.).

The diagnostic sessions explored these tasks only. We suspect that many of the other business rules applied by credit card providers would be equally if not more difficult for consumers to work out in the current designs.

Before and after comments

We received a lot of negative criticism of the credit card statements design, layout and content – both before and after use.

Even those participants that could use the statement designs at an optimum level were harsh in their criticisms. For example, the one participant that could use a statement at 100% had nothing positive to say:

‘It looks complicated, has too many colours, I can’t see at a glance what I owe. ... it even leaves a somewhat seedy impression. There is more advertising than billing information!’

One worrying aspect of the following comments is that many participants seem oblivious to the fact that they were unable to use their sample credit card design to perform any of the tasks successfully. For example, one Portuguese participant claims:

‘Its clear. The information is quite visible and explicit.’

Despite these assertions, this participant only managed to use the credit card sample at an overall level of 25%! Another satisfied participant, in the first US sample, said confidently:

'Legible. Easy to understand. The way it's listed pointed to headlines in matrix...it's all here. Good oversight. Well done. Organized. Easy to read.'

This participant scored 17%.

The frustration and irritation suffered by consumers can be sensed from the comments they made after using the credit card statements for what should be straight-forward tasks:

Table 5: Participant comments

'It's confusing and not easy to find things. Too many boxes. There's no overall structure hierarchy. I don't know what interest applies. Where are the fees and how much are payments and interest? There's too many boxes, the amount due isn't there and the late fees aren't there. How can I pay?'

'It even leaves a somewhat seedy impression.'

'There is a lot of information I don't understand or I don't know what it is there for.'

'Terrible! It is not well explained, I don't understand the vocabulary they use and the way the sum appears on the top of the list is extremely confusing.'

'This is not for people who are not used to forms. It's a disaster.'

'You should not buy such a card. It is a useless statement. It is only there to confuse you.'

'(The layout) is very confusing, too much information, often repeated, leading to doubts.'

'Bad. There are 40 things on here and I can't find anything easily! (It does not tell you) how I can pay, when the payment due date is and what the statement period is.'

'It is in such disarray – to me it looks like a high school project.'

'Sh*t! You don't know what a lot of it means.'

'Confusing! It doesn't clarify any of my spending, it's intimidating, doesn't give me information that I'd need.'

'Not easy to read, lot of information, even though it's got bold, I automatically think it's not going to be easy to understand. I don't like tables - It is too hard to follow the lines. I can't read across them.'

'It doesn't give you what you want to see. You have to search for what you want to know. It is very unclear as to whether you have made your payment and the interest charge you would incur.'

'It is more difficult than I thought and visually overwhelming. I'm glad my spouse deals with these bills.'

'I would love to see information on how to reduce interest rate and state very clearly where customer service can be accessed—it's ridiculous!'

'It feels like a deliberate withholding of information and obfuscation, I would like to see it clearer, more up-front, visually.'

'It's really confusing, made worse by the fact that I use a statement like this all the time.'

'This [marketing information] is extremely annoying and it should all go — you don't need to be told, when you get your bill, to "spend more money".'

'It's a bit too busy, and you do have to search for things. [The information on the reverse] is very small, very confusing. Some of it is all right, like the list of purchases. It's a bit messy; it could be cleaner and clearer, and slightly bigger in certain places ... Why would you be interested in all this information, and all this rubbishy advertising?'

'It's really confusing... that print on the back is absolutely tiny... There isn't any clear contact information.'

'Credit card statements are not nice things. (When I get my statements) I go 'Oh my god!' I have credit cards with two companies and they could make it a bit easier to understand... making the payments easier, and "if you do need help, contact...".'

The final comment sums it up :

It's a f*****g nightmare!

Conclusion

The picture to emerge from these findings is one of unacceptable systemic failure.

This is most tellingly illustrated by aggregating the data across all tasks for each of the statements tested. Not one of them gets to the acceptable target performance level of 81% (see Figure 9).

While consumers may be used to receiving poorly designed statements, and have even come to accept such communications as 'standard', this is not an excuse for credit card companies to continue this practice.

It is tempting to see this systemic failure as a symptom of conspiratorial action by credit card providers. However, to do so would require us to

ascribe a degree of wilful dissembling and deliberate engineering or design of the documents to make them unusable. This would require the credit card providers to have at least some skills in sophisticated information design, and there is absolutely no evidence of this in the designs we tested. Indeed, these documents look like many others to emerge out of contemporary information factories; not through a process of deliberate design, but as the end product of amateur typography and a lack of systematic and rigorous information design processes.

More likely, these documents and the pathological symptoms they display are the result of uncaring neglect. Insofar as this neglect might provide cover for some unacceptable business practices, regulators need to take firm measures to protect consumers. But based on our experience, we would advise regulators to *specify the tasks that customers should be able to perform with the documents and the acceptable level at which they should be able to do so*, leaving the execution of particular designs to professional information designers.

The current practices in some regulatory bodies is to specify the content and appearance of a document. This merely demonstrates that such regulators have the same low level of information design skills as is

revealed by industry in creating these documents. Specifying content and appearance helps no one, least of all the consumers, and any industry with predatory intentions can use their compliance with the letter of the law as a cover for predatory practices.

By specifying the tasks (and leaving with industry and professional information designers the space to enable those tasks), room is left open for innovation and for market forces to provide incentives for good design.

As we have seen in other industries, businesses which are first to market with new and innovative designs can capture a significantly increased market share, and the less innovative then copy the winning designs. In the end the customer benefits.

We were disappointed that this Communication Baseline measurements study found such uniformly poor designs, and we want to encourage industry to do better in the future.

We would like to repeat this particular study in 2015, when companies have had an opportunity to see these results and learn from them, and also after they have had time to respond to some of the newer regulatory requirements for this type of document. We would like

industry to offer us their best examples for the next Communication Baseline measurements study and we hope we can at that time publish a happier set of numbers.

From a consumer's perspective, information in relation to the interest rate, the amount of interest that has been charged, and the ways to pay or avoid paying interest or other charges appear to be deliberately unclear or difficult to follow:

'(It is) more confusing (than a normal bank statement) ...particularly the first matrix of numbers, when it's due, minimum, available balance, etc. Compacted and dense, disorienting. It seems to be less information about how to make payments. Hidden charges are intentionally nasty.'

From a regulatory perspective too, this should be unacceptable.

References

1. CRI Banking website baseline measurements studies
2. CRI Medicine labelling baseline measurements data
3. Tyers, A, & Sless D 2003
Testing Financial Service Guides Financial Products,
Financial Service Guide
Report to: Investment and Financial Services Association
Canberra: Communication Research Institute of Australia
4. CRI Government and business forms case histories baseline measurements data
5. CRI Utility bills baseline measurements data
6. Consumer Medicines Information (CMI)
7. A full account of how to conduct diagnostic testing is contained in [Writing about medicines for people.](#)
8. Beime B, Lawrence M, Rieke C (2007)
The practical application of readability user tests in national and international marketing authorisation procedures
Regulatory Rapporteur – May Issue 2007 – 13.

9. Faulkner, L. (2003)
Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. Behavior Research Methods, Instruments, & Computers
2003, 35 (3) 379-383

This study has been made possible by volunteers around the world who think this work is important.

If you would like to volunteer to be an investigator in any of our forthcoming international baseline measurement studies, please [subscribe to our newsletter](#) where we will publish details of the next studies.